



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE MEDICINA

DIVISIÓN DE ESTUDIOS DE POSGRADO

HOSPITAL INFANTIL DE MÉXICO FEDERICO GÓMEZ

“Aprendizaje de destrezas quirúrgicas endoscópicas en entrenadores físicos: confiabilidad en la evaluación de tiempo y calidad de destrezas en un modelo híbrido y comparación con la obtenida en un modelo tradicional.”

TESIS DE POSGRADO

QUE PARA OBTENER EL TÍTULO DE:

CIRUJANO PEDIATRA

PRESENTA:

DRA. KARLA VANESSA GONZÁLEZ DÍAZ

ASESORES DE TESIS:

DR. JUAN DOMINGO PORRAS HERNÁNDEZ

DR. EDUARDO BRACHO BLANCHET



HOSPITAL INFANTIL de MÉXICO

FEDERICO GÓMEZ

Instituto Nacional de Salud

MÉXICO, D. F.

2007



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.



Universidad Nacional Autónoma de México

Facultad de Medicina

División de Estudios de Posgrado

Hospital Infantil de México Federico Gómez

“Aprendizaje de destrezas quirúrgicas endoscópicas en entrenadores físicos: confiabilidad en la evaluación de tiempo y calidad de destrezas en un modelo híbrido y comparación con la obtenida en un modelo tradicional.”

Tesis de posgrado
que para obtener el título de

CIRUJANO PEDIATRA

Presenta:

Dra. Karla Vanessa González Díaz

Dr. Juan Domingo Porras Hernández

Dr. Eduardo Bracho Blanchet

Asesor de Tesis



Dedicado a:

***Mis padres por su amor, dedicación, paciencia, ejemplo y sabiduría,
gracias por estar ahí siempre, un logro más.***

A mi hermano por su cariño y apoyo, siempre juntos.

A mamá Chico, gracias por guiarme desde allá.

A mis maestros que me dieron la oportunidad, gracias por todo.

A mis compañeros que juntos vivimos momentos irrepetibles.

A los niños que me enseñan a vivir, son grandes!!!

A todos aquellos que vivieron junto conmigo estos cuatro años.

Gracias....

“Aprendizaje de destrezas quirúrgicas endoscópicas en entrenadores físicos: confiabilidad en la evaluación del tiempo y calidad de destrezas en un modelo híbrido y comparación con la obtenida en un modelo tradicional”

INVESTIGADORES

Eduardo Bracho Blanchet

Juan Domingo Porras Hernández

COINVESTIGADORES

Jaime Nieto Zermeño

Ricardo M. Ordorica Flores

Pablo Lezama del Valle

Daniel Lorías-Espinoza

Arturo Minor-Martínez

Karla Vanesa González Díaz

Cristian Zalles Vidal

INSTITUCIONES PARTICIPANTES

Departamento Cirugía General
Hospital Infantil de México Federico Gómez

Departamento de Ingeniería Eléctrica, Sección Bioelectrónica
Centro de Investigación y Estudios Avanzados, Instituto Politécnico Nacional.

ÍNDICE DE CONTENIDOS

Resumen.....	6
Marco teórico o conceptual.....	8
Antecedentes.....	10
Planteamiento del problema y Justificación	12
Pregunta de investigación.....	12
Objetivos.....	13
Hipótesis.....	13
Metodología.....	14
Desarrollo de la prueba	17
Tamaño de la muestra.....	19
Consideraciones éticas.....	19
Análisis Estadístico.....	20
Resultados.....	23
Análisis.....	27
Discusión.....	30
Conclusiones.....	31
Anexos.....	32
Referencias bibliográficas.....	35

RESUMEN

TÍTULO

Aprendizaje de destrezas quirúrgicas endoscópicas en entrenadores físicos: confiabilidad en la evaluación de tiempo y calidad de destrezas en un modelo híbrido y comparación con la obtenida en un modelo tradicional.

ANTECEDENTES

La enseñanza de la cirugía endoscópica es un reto para las instituciones donde se forman cirujanos pediatras. Se han desarrollado modelos de ambientes quirúrgicos simulados que proponen ventajas pedagógicas respecto a la objetividad, reproducibilidad y factibilidad de la evaluación de las destrezas quirúrgicas del alumno.

Existen dos grandes grupos de modelos de simulación para el entrenamiento laparoscópico: el físico y el virtual. Estos tienen diferencias importantes, sin poder sustituir uno al otro. El modelo físico propone beneficios hápticos y una física real dentro del modelo, pero la valoración de los resultados es subjetiva. Existen estudios de concordancia interobservador y prueba re-prueba del modelo físico, pero se han realizado con pocas repeticiones y con una muestra pequeña.

La Sección de Bioelectrónica del Departamento de Ingeniería Eléctrica del Centro de Investigación y Estudios Avanzados del Instituto Politécnico Nacional (CINVESTAV-IPN) desarrolló un entrenador para cirugía laparoscópica que conjunta el modelo físico y una valoración objetiva por computadora de las pruebas de transferencia y corte del sistema de evaluación de destrezas quirúrgicas endoscópicas de McGill (MISTELS).

JUSTIFICACIÓN

Existe necesidad de un modelo que pueda aportar los beneficios del modelo físico tradicional y la objetividad de medición del modelo virtual. Deseamos saber la confiabilidad en la medición de destrezas empleando un modelo que combina las características más útiles del físico tradicional y del virtual.

OBJETIVOS

1. Determinar la confiabilidad en la evaluación de tiempo y calidad de destrezas quirúrgicas endoscópicas en el modelo físico híbrido.
2. Determinar la confiabilidad en la identificación de usuarios competentes y no competentes con el modelo físico mejorado y el modelo físico tradicional.
3. Identificar factores que modifiquen la confiabilidad del modelo híbrido.

HIPÓTESIS

El uso de un modelo físico que incorpora una forma de medición de tiempo y calidad de destreza independiente de observadores externos es más confiable que un modelo físico tradicional para evaluar tiempo y calidad de destrezas e identificar usuarios competentes.

MÉTODO

Comparación transversal y ciega del tiempo y calidad en las tareas de transferencia y corte del sistema MISTELS obtenida por un grupo de usuarios en el modelo físico híbrido y en el modelo físico tradicional, en una población del Hospital Infantil de México, Federico Gómez.

Se determinarán las características diagnósticas para usuarios competentes de cada modelo y la confiabilidad de su evaluación determinando la concordancia interobservador con la concordancia cruda, valor de *phi* y *tau* de Kendall.

Se registrarán todas las complicaciones y dificultades encontradas durante la realización de los estudios.

RESULTADOS

Se efectuaron 109 pruebas a 109 sujetos con 436 mediciones, todas completándose satisfactoriamente. No hubo ninguna suspensión del estudio.

Se observó que la sensibilidad es mayor en el modelo tradicional (50%), ya que éste con más frecuencia detecta los verdaderos positivos que el entrenador híbrido (VPP 50% y VPN 91%). Esto es secundario principalmente a complicaciones generadas en la identificación de los individuos competentes y no competentes en base al cuestionario inicial, así como las alteraciones generadas por el movimiento de la base de los pivotes y de la tabla del corte lo cual conlleva a un sesgo importante concluyendo en una baja precisión.

Así mismo se evidenció que la confiabilidad es baja (de cuánto) y también esta influenciada por la poca precisión al momento de ambas mediciones.

En el análisis de prueba re prueba se identificó una concordancia mayor en el modelo tradicional que en el híbrido (cuál fue el valor de concordancia cruda, *phi* y *tau* de Kendall).

CONCLUSIÓN

El desempeño del modelo híbrido no fue el esperado, por el momento, tal y como se empleo en este estudio no es un instrumento válido para distinguir individuos competentes de los no competentes, tampoco es confiable para reproducir los resultados del entrenador físico tradicional.

A futuro haciendo mejorías tanto al Hardware como al Software, este modelo podría brindar mayor grado de objetividad en las calificaciones.

MARCO TEÓRICO

El progreso de la cirugía endoscópica ha llevado a los centros educativos quirúrgicos a implementar planes de estudio que contemplen la adquisición y desarrollo de destrezas en forma sistemática como parte importante de la preparación de los nuevos cirujanos (1).

La forma tradicional de aprendizaje de la cirugía ha sido cuestionada como ineficiente; su base en la observación y en la práctica intermitente en pacientes ofrece pocas oportunidades de aplicar en forma sistemática y continua los principios del aprendizaje de los adultos y acelerar el logro de la autonomía en el residente quirúrgico.

Los ambientes simulados ofrecen la ventaja pedagógica de facilitar la creación y prueba de experiencias de aprendizaje que cumplan un triple criterio de objetividad, reproducibilidad y factibilidad para el reforzamiento del aprendizaje y seguridad para los pacientes. El valor pedagógico de estos ambientes o modelos se establece mediante la determinación de su validez de constructo y predictiva.

La validación de constructo demuestra que un modelo es capaz de diferenciar sujetos en base a su destreza. La validación predictiva, muy importante desde el punto de vista del desempeño práctico quirúrgico, determina si un individuo tiene la destreza necesaria para resolver un problema clínico-quirúrgico real.

Existen pocos estudios que analicen esto; uno de los más relevantes fue desarrollado por Seymour y cols. en el que 16 residentes quirúrgicos fueron aleatorizados para recibir entrenamiento quirúrgico convencional o el convencional más entrenamiento repetitivo en el simulador MIST-VR de realidad virtual hasta alcanzar un nivel de experto en la tarea de manipulación y diatermia (2). Se evaluaron en forma ciega los videos de una colecistectomía laparoscópica efectuada por cada residente al final del periodo de entrenamiento determinando dos expertos el número de errores cometidos durante la cirugía. Los residentes que recibieron entrenamiento en el simulador de realidad virtual cometieron entre 5 y 10 veces menos errores transoperatorios que los entrenados convencionalmente. En muchos centros de enseñanza quirúrgica estos ambientes simulados ya forman parte integral de su plan de estudios (3,4).

Los modelos de simulación pueden dividirse en dos: los virtuales y los físicos (cajas de entrenamiento). Ambos se han comparado intentando definir cuál es más adecuado para favorecer el aprendizaje, obteniendo importantes diferencias.

El modelo virtual, más caro y de difícil acceso especialmente en países en vías de desarrollo, no cuenta generalmente con una retroalimentación háptica y puede tener una inadecuada recreación de la física provocando que algunos ejercicios no aporten una adecuada ganancia de destreza (por ejemplo, en la realización de un nudo intracorpóreo o un surgete). Sin embargo, estos modelos brindan el beneficio de la objetividad en la medición del tiempo y calidad de la destreza, lo cual es ventajoso para poder desarrollar curvas de aprendizaje reproducibles y comparables (3,5). Debido a esto se pueden utilizar para evaluar la destreza adquirida por un cirujano (2).

El modelo físico es más barato, de más fácil acceso y portátil, se utiliza instrumental laparoscópico real, tiene una sensibilidad real (retroalimentación háptica) y la física de los cuerpos dentro del modelo es la real (elemento muy importante para la práctica de nudos intracorpóreos y sutura), pero sus pruebas normalmente son más simples, sin poderse modificar tantas variables dentro del mismo ejercicio como en el virtual, y se requiere un suministro continuo de pinzas, sutura, etc. La principal desventaja es la menor objetividad en la medición de la calidad de la destreza y del tiempo (1).

Existen estudios que determinan que los modelos virtuales son más efectivos para la transferencia de una destreza a un entorno quirúrgico real. Sin embargo, por las ventajas antes mencionadas, los modelos físicos son los recomendados por la Sociedad Americana de Cirugía Gastrointestinal y Endoscópica (SAGES) para el mantenimiento y desarrollo de estas destrezas (6,7).

ANTECEDENTES

Tradicionalmente, el aprendizaje y práctica de la cirugía se ha realizado en los quirófanos mediante la interacción del experto con el aprendiz. Este modelo, postulado desde el siglo XIX por Bernhard von Langenbeck (8), se ha mantenido en gran parte del mundo hasta el momento actual. Sin embargo, se empiezan a utilizar otros métodos que incorporan algunas de las teorías educativas contemporáneas para facilitar el aprendizaje. En múltiples ramas de la medicina se han empezado a utilizar los modelos de simulación de la realidad para el desarrollo de destrezas. En un estudio prospectivo aleatorizado y ciego de 24 residentes de anestesiología, Naik y cols. probaron la superioridad para la transferencia de la destreza de intubación endotraqueal guiada endoscópicamente en pacientes reales desarrollada en un simulador físico de baja fidelidad comparada con un método teórico de aprendizaje (9).

En laparoscopia estos modelos han presentado un auge y existen trabajos que han demostrado su utilidad. En uno de los estudios más reveladores del impacto de la simulación en realidad virtual, Youngblood y cols. aleatorizaron a 46 estudiantes norteamericanos de premedicina sin experiencia previa en cirugía endoscópica para ser entrenados por 4 sesiones de 45 minutos en el simulador de realidad virtual LapSim, en una caja física de entrenamiento laparoscópico, o no recibir entrenamiento en simulador alguno, determinando en forma ciega el nivel de destreza en tareas quirúrgicas específicas en un modelo animal vivo 2 semanas después. Los estudiantes entrenados en el simulador de realidad virtual demostraron tener hasta 1.5 veces mejor desempeño que los entrenados en el simulador físico y que los controles, acercándose más dicho desempeño al efectuado por expertos que los otros grupos (7). Aunque este estudio y el de Seymour y cols. (2) favorecen la realidad virtual como modelo de aprendizaje de destrezas quirúrgicas endoscópicas, no lo hacen de manera contundente e indiscutible. El estudio de Youngblood muestra que el entrenamiento en simuladores físicos es efectivo para el desarrollo de estas destrezas (7).

En Canadá, Fried y cols. validaron el Sistema Inanimado de McGill para el Entrenamiento y Evaluación de Destrezas Laparoscópicas (MISTELS) empleando simuladores físicos tradicionales (10). Este sistema es el recomendado para el desarrollo y mantenimiento de las destrezas por el Comité de Fundamentos para la Cirugía Laparoscópica (FLS) de la SAGES (11).

La gran desventaja de la tecnología física tradicional es la subjetividad en la evaluación de la calidad de la destreza y del tiempo, dado que por las características del entrenador se requiere de un evaluador externo o de que el mismo practicante sea quien evalúe su propia destreza.

Vassiliou y cols. han determinado la confiabilidad interobservador y la confiabilidad prueba re-prueba del sistema MISTELS en un entrenador físico tradicional, concluyendo que la confiabilidad es adecuada para utilizarse en programas de entrenamiento de cirugía laparoscópica (12). Sin embargo, esta determinación se basa en resultados obtenidos en una muestra de 12 personas realizando 2 veces la prueba con dos observadores entrenados. En total, cada observador evaluó sólo 24 repeticiones. Éstos pueden ser defectos metodológicos importantes debido a que las condiciones de este estudio no reflejan la realidad de una experiencia educativa en cirugía endoscópica. Se sabe que se necesitan realizar de 10 a 30 repeticiones de cada ejercicio para obtener un aprendizaje adecuado y de dos a cinco para determinar la habilidad basal de una persona (13). En un ambiente simulado físico un observador externo está expuesto a una mayor probabilidad de error y de cansancio, siendo posible que las determinaciones de tiempo y precisión no sean fidedignas después de 30 repeticiones, aun cuando se trate de un solo usuario.

PLANTEAMIENTO Y JUSTIFICACIÓN DEL PROBLEMA

Existe la necesidad de un entrenador que facilite el aprendizaje de las destrezas quirúrgicas endoscópicas que sea de fácil acceso para los cirujanos, que facilite una práctica continua para mantener la destreza y que cuente con una forma objetiva y reproducible, independiente del observador externo o del practicante, en la medición del tiempo y la evaluación de la calidad de la destreza.

En la Sección de Bioelectrónica del Departamento de Ingeniería Eléctrica del Centro de Investigación y Estudios Avanzados del Instituto Politécnico Nacional (CINVESTAV-IPN) se desarrolló un entrenador para cirugía laparoscópica que, en conjunto con una computadora de escritorio o portátil, evalúa en forma independiente el tiempo y la calidad de las destrezas de transferencia y corte del sistema MISTELS. La mejora del simulador físico tradicional con una forma de objetividad en la evaluación de la destreza que se pueda mantener en el mismo nivel de confiabilidad independientemente del número de repeticiones, momento del día, tipo y número de usuarios puede representar una ventaja para construir secuencias de aprendizaje con resultados confiables y reproducibles.

Demostrar su validez de constructo para distinguir a un usuario competente en destrezas quirúrgicas endoscópicas (14) y su confiabilidad en la evaluación del tiempo y calidad de dichas destrezas comparándola con las del entrenador tradicional puede representar un avance en la búsqueda de una tecnología educativa versátil para mejorar la calidad de las experiencias de aprendizaje para los cirujanos fuera de los quirófanos.

PREGUNTA DE INVESTIGACIÓN

¿Cuál es la confiabilidad de la evaluación de destrezas quirúrgicas endoscópicas en un modelo físico que incorpora una forma de medición de tiempo y calidad de destreza independiente de observadores externos?

¿Qué factores, especialmente en la medición del tiempo y precisión, podrían afectar la confiabilidad del modelo híbrido?

OBJETIVOS

1. Determinar la confiabilidad en la identificación de usuarios competentes y no competentes con el modelo físico mejorado y el entrenador físico tradicional.
2. Determinar la confiabilidad en la evaluación de tiempo y calidad de destrezas quirúrgicas endoscópicas en el modelo físico mejorado desarrollado por el CINVESTAV-IPN.
3. Identificar elementos que pueden modificar la medición del tiempo y la precisión durante el desarrollo de las pruebas de corte y transferencia del MISTELS en los modelos híbrido y tradicional.

HIPÓTESIS

El uso de un modelo físico que incorpora una forma de medición de tiempo y calidad de destreza independiente de observadores externos es más confiable que un modelo físico tradicional para evaluar tiempo y calidad de destrezas e identificar usuarios competentes.

METODOLOGÍA

A. Diseño.

Comparación transversal y ciega del tiempo y calidad en las tareas de transferencia y corte del sistema MISTELS obtenida por un grupo de usuarios en el modelo físico mejorado y en el modelo físico tradicional.

B. Población de estudio.

Siete cirujanos pediatras adscritos al Departamento de Cirugía General del Hospital Infantil de México Federico Gómez (HIMFG), ocho residentes de segundo a cuarto año de la especialidad de cirugía pediátrica, cuatro residentes de primer año de la especialidad de cirugía pediátrica y 90 residentes de primero a tercer años de la especialidad de pediatría del programa único de especialidades médicas de la UNAM con sede en el HIMFG, registrados como alumnos activos de dicho programa y sede entre marzo y agosto del 2007.

C. Criterios de inclusión.

- Cirujanos pediatras adscritos al Departamento de Cirugía General del HIMFG con más de 70 procedimientos laparoscópicos como cirujanos.
- Residentes de segundo a cuarto años del programa de la especialidad en cirugía pediátrica con sede en el HIMFG con más de 5 procedimientos laparoscópicos como cirujanos o ayudantes.
- Residentes de primer año del programa de la especialidad en cirugía pediátrica con sede en el HIMFG con menos de 5 procedimientos laparoscópicos como ayudantes.
- Residentes del programa de la especialidad en pediatría médica con sede en el HIMFG sin contacto previo alguno como ayudante en procedimiento laparoscópico alguno.

D. Criterios de Exclusión.

Adscrito o residente que se niegue a para participar en el estudio.

E. Criterios de Eliminación.

Sujetos que no concluyan todas las pruebas en ambos modelos como se especifica en el protocolo.

F. Definición Operacional de Variables.

VARIABLES INDEPENDIENTES:

- a) **Modelo físico híbrido:** Modelo físico en el cual la valoración del tiempo, precisión y calidad de las pruebas se efectúa mediante el programa realizado por el CINVESTAV – IPN. El tiempo mediante el marcaje desde que entran las pinzas en el campo de visión del modelo hasta el momento en que salen posterior a terminar la prueba. En la prueba de transferencia, la precisión se determina en base al número de piezas transferidas al final de la prueba y en la de corte en base al área de la figura recortada.

- b) **Modelo físico tradicional:** Modelo físico en el cual la valoración del tiempo, precisión y calidad, se efectúan mediante la determinación de un observador externo o del mismo ejecutante de las pruebas. Se determinara el tiempo mediante la medición con un cronometro, desde el momento en que se observa que las pinzas entran en la imagen grabada hasta el momento en que las pinzas salen de la imagen, después de haber terminado el ejercicio de corte o transferencia. La precisión en la prueba de transferencia se efectúa en base al número de piezas transferidas al final de la prueba y en la de corte en base al área calculada de la figura recortada.

Clasificación de los usuarios en base a la experiencia en cirugía endoscópica:

- 1. Usuario con destrezas básicas:** Sujeto sin entrenamiento previo alguno ni experiencia previa alguna en cirugía laparoscópica.
Parámetro: Sí, No.
- 2. Usuario con destrezas intermedias:** Sujeto que realice cirugía laparoscópica de forma regular como ayudante, que cuente con menos de 10 procedimientos como cirujano o que haya recibido algún programa de entrenamiento en laparoscopia.
Parámetro: Sí, No.
- 3. Usuario con destrezas avanzadas:** Sujeto que realice cirugía laparoscópica de forma regular como cirujano, que cuente con más de 70 procedimientos laparoscópicos como cirujano o ayudante.
Parámetro: Sí, No.
- 4. Usuario competente:** Sujeto que realiza cirugía laparoscópica de forma regular como ayudante o cirujano, con 10 o más procedimientos quirúrgicos endoscópicos como cirujano o ayudante, o que ha cubierto más de 7 horas de entrenamiento supervisado en un entrenador físico y 6 horas de entrenamiento supervisado en entrenador de realidad virtual.
Parámetro: Sí, No.
- 5. Usuario no competente:** Sujeto que no ha realizado cirugía laparoscópica como ayudante o cirujano, y que no ha recibido curso alguno de entrenamiento laparoscópico.
Parámetro: Sí, No.

Variables dependientes:

- a) **Tiempo:** Duración de la tarea de transferencia o corte del sistema MISTELS, contada a partir de la introducción de las pinzas o tijeras de laparoscopia al entrenador hasta el retiro de las mismas al haber concluido la tarea. En el entrenador físico tradicional lo contará un observador externo con cronómetro digital, ciego al usuario y con acceso exclusivo a la imagen de la pantalla del televisor donde se efectúa la tarea. En el entrenador físico mejorado lo contará una computadora conectada al modelo.
Parámetro: Segundos.
- b) **Precisión en la tarea de transferencia:** Número de piezas no transferidas o colocadas fuera del campo de visión cada vez que se repita esta tarea del sistema MISTELS (10).
Parámetro: Porcentaje.
- c) **Calidad de la tarea de transferencia:** Integración del tiempo empleado por cada usuario para completar la tarea y su precisión de acuerdo con la fórmula: $[300-(\text{tiempo}+\text{precisión})]/3$.
Parámetro: Unidades.
- d) **Precisión en la tarea de corte:** Diferencia en el área calculada del círculo recortado y el área del círculo modelo de 4.5cm de diámetro (15.9 cm^2) del sistema MISTELS (10).
Parámetro: Porcentaje.
- e) **Calidad de la tarea de corte:** Integración del tiempo empleado por cada usuario para completar la tarea y su precisión de acuerdo con la fórmula: $300-(\text{tiempo}^2+\text{precisión})$.
Parámetro: Unidades.
- f) **Desempeño de avanzado en tarea de transferencia:** Promedio de la calidad de 2 repeticiones de la tarea de transferencia, mayor o igual a 70 (10).
Parámetro: Sí, No.
- g) **Desempeño de intermedio en tarea de transferencia:** Promedio de la calidad de 2 repeticiones de la tarea de transferencia, mayor o igual a 60 y menor o igual a 69 (10).
Parámetro: Sí, No.
- h) **Desempeño de básico en tarea de transferencia:** Promedio de la calidad de 2 repeticiones de la tarea de transferencia, menor o igual a 59 (10).
Parámetro: Sí, No.

- i) **Desempeño de avanzado en tarea de corte:** Promedio de la calidad de 2 repeticiones de la tarea de corte, mayor o igual a 56 (10).
Parámetro: Sí, No.
- j) **Desempeño de intermedio en tarea de corte:** Promedio de la calidad de 2 repeticiones de la tarea de corte, mayor o igual a 45 y menor o igual a 55 (10).
Parámetro: Sí, No.
- k) **Desempeño de básico en tarea de corte:** Promedio de la calidad de 2 repeticiones de la tarea de corte, menor o igual a 44 (10).
Parámetro: Sí, No.

DESARROLLO DE LA PRUEBA

Se invitará a miembros del cuerpo médico del HIMFG a participar, incluyendo médicos adscritos al servicio de cirugía pediátrica, residentes de cirugía pediátrica, residentes de pediatría, subespecialistas e internos. Cada persona que acepte participar en el estudio llenará un cuestionario y firmará una hoja de consentimiento informado (Anexo 1 y 2). Mediante el cuestionario se clasificará dentro del grupo de avanzados, intermedios o básicos y competentes o no competentes, de acuerdo a las definiciones operacionales arriba descritas.

A cada individuo que ingrese al estudio se le asignará un número para su identificación durante la realización del protocolo.

Todos los participantes del estudio recibirán una explicación estandarizada y por la misma persona del objetivo del estudio, se les proyectará un video de cómo se efectúa cada tarea y contarán con dos minutos previos al desarrollo de la prueba para familiarizarse con el instrumental, la imagen del monitor, la interacción con el entrenador y el material a emplear. Cualquier duda será resuelta de inmediato en forma verbal por el mismo inductor.

Se contará con el modelo de entrenamiento laparoscópico, ubicado en un sitio cerrado fijo, en donde contaremos con el material necesario para el desarrollo de las pruebas. Una persona encargada del adecuado uso del modelo y que pueda resolver cualquier problema que se suscite se encontrará presente durante la realización de las pruebas.

Según vayan acudiendo a realizar las pruebas se realizarán los siguientes pasos:

- 1) Corroborar que tenga lleno el cuestionario inicial. En caso de requerirse, se completará en ese momento.
- 2) Asignar un número aleatorio determinado por una tabla de números aleatorios al momento de su llegada. Este número se anotará en su

hoja inicial, con el cual se le identificará durante el desarrollo del estudio, la recolección de resultados y análisis de datos. Nunca se utilizará otro elemento para identificarse, esto para permitir una valoración ciega de los resultados.

- 3) Permitir que la persona pruebe el modelo, con la prueba de transferencia y corte. Se resolverán las dudas que pudieran originarse.
- 4) El participante realizará dos repeticiones de cada prueba. Todo el ejercicio será grabado en video. Cada participante se identificará en el video con su número el cual se grabará al inicio de la prueba. Las pruebas se efectuarán de lunes a viernes de 8 a 16:00hrs. Se empleará como monitor un televisor a color con pantalla de 15 pulgadas colocado a la altura de los ojos del alumno y a una distancia de 1.5m. El instrumental a utilizar será una pinza endoscópica de disección de 5mm desechable (Endo Dissect™5mm, United Status Surgical, Tyco Healthcare), una pinza endoscópica de sujeción de 5mm desechable (Endo Clinch II™5mm, United States Surgical, Tyco Healthcare) y una tijera curva endoscópica de 5mm desechable (Endo Shears™5mm, United States Surgical, Tyco Healthcare). Para el modelo mejorado, se empleará una computadora portátil con procesador Intel-Pentium 4, de 1.2GHz, y con sistema Windows XP, a la cual se le instaló previamente el software diseñado por el laboratorio de bioelectrónica del CINVESTAV-IPN.
- 5) La valoración mediante el modelo mejorado se efectuará en tiempo real durante la realización de cada tarea. El cálculo del tiempo lo hará automáticamente el programa desarrollado por el CINVESTAV-IPN, al igual que cálculo del área del círculo de tela resultante de la prueba de corte. Estos resultados se anotarán en la hoja de recolección de datos.
- 6) La valoración clásica se realizará mediante mediciones del tiempo y número de piezas transferidas realizadas a los videos grabados y a de la pieza de tela resultante de la prueba de corte. Esta valoración se efectuará en un momento posterior a las pruebas, así nos aseguraremos que la persona que realice las mediciones no pueda estar en contacto con los individuos que efectúen las pruebas. Para estas mediciones se utilizará un cronómetro con centésimas de segundo (DT3 Digi Sport Instruments, Biarritz, Francia) y se medirá el diámetro máximo del fragmento de tela recortado en la tarea de corte empleando un flexómetro con escala milimétrica de aluminio (FR3W, Wilson, México). Se anotarán los resultados en hojas de recolección de datos. En las tareas de transferencia y corte el tiempo empieza a correr al momento en que aparece en pantalla por primera vez alguna de las pinzas y termina al transferir la última pieza o desprender el círculo recortado. La precisión de cada tarea se determinará de acuerdo con los criterios descritos en el anexo 2.

TAMAÑO DE LA MUESTRA

Según algunos estudios, existen actualmente en los Estados Unidos de América de 17 000 a 23 000 cirujanos generales (15). La SAGES, organización que agrupa a todos los cirujanos norteamericanos que efectúan cirugías endoscópicas como parte integral del manejo de sus pacientes, tiene un registro de 5000 miembros activos en el 2006 (16). Podemos suponer que de toda la población de cirujanos generales norteamericanos, un 25% practica cirugía endoscópica con un nivel de competencia suficiente para que sea parte de su repertorio terapéutico (existen miembros extranjeros de SAGES por lo cual el porcentaje puede ser menor). En México no tenemos datos adecuados para calcular la prevalencia de cirujanos competentes en cirugía endoscópica, pero suponemos que el porcentaje es menor, probablemente un 15%.

Para el presente estudio, el tamaño de la muestra se calculó considerando los siguientes criterios: para una prevalencia de usuarios competentes del 15%, una sensibilidad promedio de las pruebas de destreza para identificar usuarios competentes del 90% y un límite inferior del intervalo de confianza para dicha sensibilidad mínimo del 55%, el número calculado de usuarios competentes a incluir en el estudio es de 18 y de no competentes de 93, de acuerdo con las tablas y fórmula de Flauhault y cols (17).

CONSIDERACIONES ÉTICAS

El presente estudio está de acuerdo con los principios éticos para las investigaciones médicas en seres humanos promulgados por la Asociación Médica Mundial en la Declaración de Helsinki, enmendados en la 52ª Asamblea General de Edimburgo, Escocia, en octubre del 2000. Se explicará detalladamente al participante en qué consiste el estudio, sus beneficios, riesgos y costos. El equipo multidisciplinario se compromete a garantizar el manejo confidencial de la información clínica de cada participante. No se realizará ningún tipo de procedimiento o práctica en pacientes u organismo vivo. El estudio deberá ser aprobado por los comités de investigación y ética del HIMFG.

ANÁLISIS ESTADÍSTICO

1. Comportamiento de las tareas de transferencia y corte como pruebas para identificar usuarios competentes.

Para cada punto de corte empleado en las destrezas se calcularán sus propiedades diagnósticas en el entrenador físico tradicional y en el mejorado, construyendo para cada entrenador y destreza el siguiente cuadro de doble entrada:

		Nivel de destreza (según respuestas en cuestionario)		Totales
		Competente	No competente	
Resultado de la prueba en el entrenador	Competente	a (verdaderos positivos)	b (falsos positivos)	a+b (total de positivos de la prueba)
	No competente	c (falsos negativos)	d (verdaderos negativos)	c+d (total de negativos de la prueba)
Totales		a+c (total de competentes)	b+d (total de no competentes)	a+b+c+d (total de usuarios estudiados)

Definiciones y cálculos de las propiedades diagnósticas de los entrenadores:

- **Sensibilidad (Sn):** Proporción de usuarios competentes de acuerdo con las respuestas obtenidas en el cuestionario inicial que tiene un resultado positivo de la prueba en el entrenador (verdaderos positivos/total de usuarios competentes ó $a/a+c$).
- **Especificidad (Sp):** Proporción de usuarios no competentes de acuerdo con las respuestas obtenidas en el cuestionario inicial que tiene un resultado negativo de la prueba en el entrenador (verdaderos negativos/total de usuarios no competentes ó $d/b+d$).
- **Valor predictivo positivo (VP+):** Proporción de usuarios con un resultado positivo de la prueba que son competentes (verdaderos positivos/total de positivos de la prueba ó $a/a+b$).
- **Valor predictivo negativo (VP-):** Proporción de usuarios con un resultado negativo de la prueba que no son competentes (verdaderos negativos/total de negativos de la prueba ó $d/c+d$).
- **Razón de probabilidad positiva (LR+):** Razón de la probabilidad de tener un resultado positivo de la prueba en usuarios competentes entre la probabilidad de tener un resultado positivo de la prueba en usuarios no competentes. Esto es: [(verdaderos positivos/total de competentes) / (falsos positivos/total de no competentes)] ó $[(a/a+c) / (b/b+d)]$. Como $a/a+c$ es igual a la sensibilidad, y $b/b+d$ es igual a $1-(d/b+d)$, siendo $d/b+d$ igual a la especificidad, esta última fórmula se puede reescribir como: $LR+ = \text{sensibilidad}/(1-\text{especificidad})$.
- **Razón de probabilidad negativa (LR-):** Razón de la probabilidad de tener un resultado negativo de la prueba en usuarios competentes entre la probabilidad de tener un resultado negativo de la prueba en usuarios no competentes. Esto es: [(falsos negativos/total de competentes) / (verdaderos negativos/total de no competentes)] ó $[(c/a+c) / (d/b+d)]$. Como $c/a+c$ es igual a $1-(a/a+c)$ siendo $a/a+c$ la sensibilidad, y siendo $d/b+d$ igual a la especificidad, esta última fórmula se puede reescribir como: $LR- = (1-\text{sensibilidad})/\text{especificidad}$.
- **Probabilidad pre-prueba de usuarios competentes o prevalencia:** la proporción del total de usuarios estudiados que en verdad son competentes. Esto es: total de usuarios competentes/total de usuarios estudiados ó $(a+c)/(a+b+c+d)$.

2. Determinación de la confiabilidad en la evaluación de tiempo y calidad de la destreza, y en la identificación de usuarios competentes y no competentes.

a) Confiabilidad interobservador.

Se determinará la confiabilidad cruda en las calificaciones de tiempo y de calidad de la destreza en cada entrenador de acuerdo con la siguiente fórmula:

Confiabilidad cruda = (Número de calificaciones en el entrenador físico y en el entrenador mejorado que concuerdan) / Número total de calificaciones.

Se interpretará como confiabilidad cruda significativa aquella mayor o igual a 0.9.

Se determinará la concordancia independiente del azar en la identificación de usuarios competentes y no competentes entre el entrenador tradicional y el mejorado con el cálculo de *phi* (18). Se construirá un cuadro de doble entrada para los usuarios competentes y no competentes con cada entrenador:

		Entrenador tradicional	
		Competente	No competente
Entrenador mejorado	Competente	a (verdaderos positivos)	b (falsos positivos)
	No competente	c (falsos negativos)	d (verdaderos negativos)

Se calculará el valor de *phi* de acuerdo con la siguiente fórmula (19):

$$Phi = [\sqrt{(ad/bc) - 1}] / [\sqrt{(ad/bc) + 1}]$$

La concordancia independiente del azar para identificar usuarios competentes y no competentes se interpretará como pobre con un valor de *phi* ≤0.4, moderado >0.4 y ≤0.6, sustancial >0.6 y ≤0.8, y significativo >0.8 (19).

Se determinará la concordancia en el valor de la calidad de cada tarea otorgado con cada entrenador en cada evaluación de los usuarios calculando el coeficiente *tau* de Kendall de acuerdo con la siguiente fórmula (20):

$$Tau = (n_c - n_d) / [n(n-1) / 2]$$

Donde:

n_c = número de pares concordantes

n_d = número de pares discordantes

n = número de pares

Se interpretará como concordancia significativa un valor de $tau \geq 0.8$ y < 0.9 , y concordancia ideal un valor ≥ 0.9 .

b) Confiabilidad prueba-reprueba.

Se calculará el coeficiente *alpha* de Cronbach para determinar la concordancia en los valores de la calidad de la destreza en cada entrenador al repetir una vez la tarea de transferencia, y una vez la de corte, de acuerdo con la siguiente fórmula (21):

$$Alpha = (n\check{r}) / [1 + (1-n)\check{r}]$$

Donde:

n = número de repeticiones

\check{r} = correlación promedio entre repeticiones

Se interpretará como una confiabilidad prueba-reprueba significativa un valor de *alpha* ≥ 0.8 y < 0.9 , e ideal ≥ 0.9 .

RESULTADOS

Se efectuaron 109 pruebas a 109 sujetos con 436 mediciones, todas completándose satisfactoriamente. No hubo ninguna suspensión del estudio.

El tiempo para la realización de cada prueba varió de acuerdo a cada sujeto entre 10 y 30 minutos.

Hubo diferencias en cuanto a la capacidad de entendimiento en la realización de las pruebas; evidentemente en sujetos “no competentes” hubo necesidad de una explicación más amplia y un mayor adiestramiento en el uso del instrumental. Es importante mencionar que a pesar de esto todos lograron entender en su totalidad las indicaciones y la realización de las pruebas.

Se tiene que coordinar de forma adecuada el momento en el que se introducen las pinzas al campo de visión y el momento en el que se inicia el conteo del tiempo con el cronómetro ya que esto puede generar diferencias importantes en la contabilidad del mismo.

A algunos participantes se les dificultó realizar la prueba de transferencia debido a que el tablero en donde se encuentran los pivotes se desplazaba y esto generaba alteraciones en la visualización y orientación de las piezas y del tablero, por ende dificultad en concluir con éxito la prueba (22/109). Esto influye de manera importante ya que el sistema híbrido es ineficaz para contabilizar de manera adecuada las piezas realmente transferidas al momento de moverse el tablero. Así mismo observamos que las piezas transferidas deben quedar idealmente separadas del pivote ya que éstas, al estar en contacto con él, no son calificadas adecuadamente como sucedió en 63 de nuestros sujetos evaluados.

En la prueba de corte sucedió lo mismo aunque con menos frecuencia, sin embargo a muchos participantes les parecía que la tela en la que se realizaba el corte quedaba flácida y esto les generaba retardos en la terminación de la prueba. Generalmente se intentó terminar el corte en los minutos establecidos a pesar de la precisión.

La confiabilidad cruda para la prueba de transferencia fue de 0.87 y la de corte es de 0.80. El valor de phi para corte es de 0.92 mientras que para la transferencia no se realizó debido a que el resultado de cero en los falsos positivos no permite realizar la fórmula.

El valor de tau para la transferencia es de 0.88 y para corte de 0.87.

El valor kappa para corte es de 0.4 y para transferencia de 0.75.

Comportamiento de la prueba de transferencia en el modelo híbrido para identificar usuarios competentes

	Competente	No Competente	
Competente	0	1	1
No Competente	16	92	108
	16	93	109

Sensibilidad	0
Especificidad	98%
Valor Predictivo Positivo	0
Valor Predictivo Negativo	85%
Razón de Probabilidad Negativo	1
Razón de Probabilidad Positivo	0
Prevalencia	14.6%

Comportamiento de la prueba de transferencia en el modelo tradicional para identificar usuarios competentes

	Competente	No Competente	
Competente	8	8	16
No Competente	8	85	93
	16	93	109

Sensibilidad	50%
Especificidad	91%
Valor Predictivo Positivo	50%
Valor Predictivo Negativo	91%
Razón de Probabilidad Negativo	54%
Razón de Probabilidad Positivo	6.25%
Prevalencia	14.6%

Comportamiento de la prueba de corte en el modelo híbrido para identificar usuarios competentes

	Competente	No Competente	
Competente	6	10	16
No Competente	10	83	93
	16	93	109

Sensibilidad	37%
Especificidad	89%
Valor Predictivo Positivo	37%
Valor Predictivo Negativo	89%
Razón de Probabilidad Negativo	0.69%
Razón de Probabilidad Positivo	3.7%
Prevalencia	14.6%

Comportamiento de la prueba de corte en el modelo tradicional para identificar usuarios competentes

	Competente	No Competente	
Competente	15	15	30
No Competente	1	78	79
	16	93	109

Sensibilidad	93%
Especificidad	83%
Valor Predictivo Positivo	50%
Valor Predictivo Negativo	98%
Razón de Probabilidad Negativo	0.07%
Razón de Probabilidad Positivo	5.8%
Prevalencia	14.6%

Cuadro de comparación modelo tradicional contra modelo híbrido en la prueba de transferencia

		<i>Modelo tradicional</i>		
		<i>Competente</i>	<i>No Competente</i>	
<i>Modelo híbrido</i>	<i>Competente</i>	1	0	1
	<i>No Competente</i>	15	93	108
		16	93	109

Valor Tau = 0.88

Valor Phi = indeterminado

Concordancia cruda = 0.87

Cuadro de comparación modelo tradicional contra modelo híbrido en la prueba de corte

		<i>Modelo tradicional</i>		
		<i>Competente</i>	<i>No Competente</i>	
<i>Modelo híbrido</i>	<i>Competente</i>	12	4	16
	<i>No Competente</i>	18	75	93
		30	79	109

Valor de Tau = 0.87

Valor de Phi = 0.92

Concordancia cruda = 0.80

TABLA DE COMPARACIÓN DE TIEMPOS ENTRE AMBOS MODELOS EN LA PRUEBA DE TRANSFERENCIA

Modelo Híbrido	236 seg / persona	Valor de p
Modelo Tradicional	223 seg / persona	< 0.001

TABLA DE COMPARACIÓN DE TIEMPOS ENTRE AMBOS MODELOS EN LA PRUEBA DE CORTE

Modelo Híbrido	250 seg / persona	Valor de P
Modelo Tradicional	237 seg / persona	0.002

ANÁLISIS

Para el análisis adecuado de nuestros resultados, dividiremos esta sección en 2 partes, una que trata la prueba de transferencia y otra la de corte, ya que cada ejercicio tiene diferentes cualidades.

En el modelo híbrido, la prueba de transferencia, muestra una baja capacidad para detectar a los individuos competentes y secundariamente presenta una muy buena capacidad para detectar a los no competentes. La prueba en modelo híbrido sólo clasificó a un individuo como competente siendo que éste no lo era en base a los datos obtenidos del cuestionario. Esta diferencia es atribuible al cálculo de la precisión, el entrenador híbrido calculó una precisión de 6.8 en 2 repeticiones comparado con la de 9.8 en 2 repeticiones del entrenador tradicional, con una diferencia de 30.9%. En cambio, el tiempo calculado por los 2 modelos resultó similar, con una diferencia únicamente de 5.2%. Esta diferencia en el cálculo de precisión puede deberse a la forma en que el modelo híbrido cuantifica el número de piezas transferidas:

a) el programa no cuenta una pieza como transferida si está muy cerca del pivote donde se transfiere, y si está centrada sin tocar el pivote cuenta como correcta.

b) Si la base de los pivotes se mueve durante el estudio (la mayor parte de las veces secundario a uso excesivo de fuerza por parte del practicante) las piezas contabilizadas difieren del real. El uso de fuerza excesiva es un factor importante que en muchos modelos virtuales se cuenta para calcular la habilidad del practicante. Sin embargo, no es el objetivo en esta prueba, no se ha establecido su utilidad predictiva y en el modelo híbrido si creó un sesgo de medición.

Llama la atención que los resultados obtenidos con el evaluador externo no fueron los esperados idealmente, ya que la sensibilidad fue sólo del 50% y la especificidad de 91%. Esto es atribuible a algunas situaciones hipotéticas como falta de tiempo para acostumbrarse al uso del modelo (ejercicio, pinzas etc), movilización de la base de la prueba al realizar los ejercicios y una mala clasificación de los individuos como competentes mediante el cuestionario.

En la prueba de corte, la sensibilidad para detectar individuos competentes con el modelo híbrido fue de 37% comparado con un 93% del modelo tradicional. La especificidad fue mas eficientemente calculada por el entrenador híbrido que por el evaluador externo, igualmente el valor predictivo negativo fue mejor mediante el modelo híbrido. También en esta prueba la principal dificultad del modelo es determinar a los individuos competentes como competentes (verdaderos positivos), pudiendo atribuir también esto a que varios individuos competentes tardaron en acostumbrarse al uso del modelo, lo cual pudo afectar su desempeño en la prueba. El tiempo calculado por los dos modelos fue similar, con una diferencia de 3.7%, por lo cual la principal complicación fue la determinación de la precisión que al ser resultado del cálculo del área, coloca en desventaja directa al modelo tradicional, ya que no es prácticamente posible calcular una área de una figura irregular, en cambio es un calculo fácilmente realizado por el programa de computación.

Estos problemas del modelo híbrido condicionaron imprecisión en la discriminación de individuos competentes los no competentes. Esto, a su vez, se manifiesta en las bajas concordancias cruda e independiente del azar entre los modelos para identificar usuarios competentes y no competentes:

La confiabilidad cruda entre ambos modelos (Interobservador), fue adecuada, ya que ambos se encuentran debajo del 0.9 como limite, con una confiabilidad (concordancia) de 0.87 para la prueba de transferencia y de 0.80 para la prueba de corte. El resultado de la concordancia independiente del azar, calculado mediante la prueba de phi, por la ausencia de falsos negativos, se obtiene un valor infinito. En la prueba de corte, la confiabilidad fue de 0.92, la cual es significativa.

La concordancia para la prueba de corte fue pobre con un valor de phi de 0.92. Mientras que la prueba de transferencia no se logró realizar ya que no se detectaron falsos positivos, lo cual representa una concordancia no significativa entre ambos modelos. Estos valores son esperados, debido a que el modelo híbrido tuvo dificultades para discriminar a individuos competentes de los no competentes.

La concordancia de los modelos en la evaluación de la calidad de cada tarea, tuvo un valor de tau significativo de 0.87 y 0.88 para corte y transferencia respectivamente.

En la confiabilidad prueba-reprueba, el coeficiente alpha de Cronbach en la transferencia tuvo un valor de 0.96 y en la prueba de corte 0.95, lo cual refleja que si hay diferencia significativa entre ambas pruebas y que tienen un coeficiente de coorelación positivo. Considerando el porcentaje de concordancia cruda, el modelo híbrido tiene una concordancia de 70% y el modelo tradicional de 75%. En la prueba de corte el coeficiente resulta de 6.69% en el modelo híbrido y de 98% en el modelo tradicional, siendo la concordancia ideal para el modelo tradicional y malo para los demás. Las variaciones que existen se pueden atribuir a las diferencias en el desempeño de cada individuo entre la primera y segunda repeticiones. Esto puede reflejar el inicio de una curva de aprendizaje, en la cual cada usuario mejora significativamente su desempeño en la segunda repetición en base a los errores que tuvo en la primera.

DISCUSIÓN

El desempeño del modelo híbrido no fue el esperado, en mayor parte por los problemas que se presentaron al realizar las pruebas (interacción de la persona con el modelo), afectando esto de forma importante el cálculo de la precisión.

El modelo híbrido brinda una objetividad que no se puede obtener con el modelo tradicional, pero este modelo todavía necesita modificaciones, tanto en el hardware como el software, para poder sustituir al modelo tradicional.

El hardware necesita mejorar en el diseño de la base con pivotes empleada en la tarea de transferencia para evitar su movimiento durante la toma de las piezas, sin que salga del área de foco de la cámara electrónica que transmitirá la información a la computadora para dar la calificación al final de la prueba. Esta alteración produce un menor número de piezas contabilizadas como transferidas. Este problema no fue tan relevante en la tarea de corte, ya que el movimiento no es parte de la calificación.

El Software en ocasiones no reconoció el número de piezas realmente transferidas por estar en contacto la rondana con la base del pivote, subestimando con esto la precisión. En la prueba de corte, no contamos con tales errores siendo la cuantificación de la precisión exacta debido a que se realiza en base a un área la cual no se modifica.

Por todo esto, podemos concluir que el modelo híbrido en su forma actual es promisorio para brindar la objetividad de la calificación de estas pruebas, pero requiere aún mayor desarrollo y mejorar el diseño de los materiales empleados para la prueba, así como mejorar los sistemas de reconocimiento óptico de la posición de cuerpos, particularmente los transferidos.

Queda pendiente definir cuántas repeticiones son necesarias efectuar en este modelo, para poder suponer que la calificación obtenida es la que verdaderamente refleje la habilidad de la persona. Con dos pruebas encontramos mucha diferencia entre la primera y segunda calificación no estando necesariamente relacionada con la habilidad de la persona.

El modelo híbrido, tal como se empleó en el presente estudio, demostró que no es un instrumento válido (es decir, que mida lo que realmente debe medir) para discriminar individuos competentes de los no competentes en cirugía endoscópica, ni confiable (que sus resultados concuerden cada vez que se emplee). Aparentemente, estos resultados pueden atribuirse a un mal diseño del algoritmo del software para evaluar la calidad de la tarea, y a un mal diseño del material para efectuar cada tarea.

CONCLUSIONES

1. El modelo híbrido, en su estado actual, no es un instrumento válido para discriminar individuos competentes de los no competentes.
2. El modelo híbrido, en su estado actual, no es un instrumento confiable para reproducir la precisión de la destreza con una repetición.
3. El modelo híbrido, en su estado actual, no es un instrumento confiable para reproducir los resultados de la precisión de la destreza obtenidos con el modelo tradicional.

ANEXO 1

No. de identificación	
------------------------------	--

Nombre:
Edad:
Sexo:

No. de cirugías laparoscópicas (cirujano)	
No. de cirugías laparoscópicas (ayudante)	
Recibió algún curso de laparoscopia (práctico)	SI cuál? _____ ¿Cuántas horas de práctica tuvo? _____ NO

Puesto (Marcar 1 opción)	Médico adscrito al servicio de cirugía.
	Médico residente de cirugía pediátrica.
	Médico residente de pediatría o médico interno.

Protocolo

ANEXO 2

MISTELS (Mc Gill Inanimate System for training and evaluation of laparoscopic skills)

- **TRANSFERENCIA**

Dos tableros y 6 ejes, tomar una pieza con mano izquierda del tablero izquierdo, transferirla en el aire a la mano derecha, y colocarla en tablero derecho.

Primero todas las fichas hacia la derecha, luego todas a la izquierda.

Objetivos: Orientación visuoespacial y ambidestreza.

Corre tiempo: Desde que entran ambas pinzas al simulador hasta terminar la tarea.

Límite de tiempo: 300 seg.

Calificación del tiempo si se excede el límite: 0

Precisión: (Número de piezas no transferidas o colocadas fuera del campo de visión/12) (100)

Cálculo de calificación: $((300 - \text{tiempo en seg}) - \text{Precisión})/3$

Discriminación: Novato: 42 (36-48)

Intermedio: 65 (60-71)

Experto: 76 (72-80)

- **CORTE**

Cortar un círculo de 4 cm de diámetro en el centro de una gasa de 10x10 cm, sostenida entre pinzas caimán.

Objetivos: Orientación visuoespacial, coreografía ambas manos.

Corre tiempo: Desde que entran ambas, pinza y tijera al simulador hasta terminar la tarea.

Límite de tiempo: 300 seg.

Calificación del tiempo si se excede el límite: 0

Precisión: Porcentaje en que difiere el círculo recortado del marcado.

Cálculo de calificación: $((300 - \text{tiempo en seg}) - \text{Precisión})$

Discriminación: Novato: 38 (34-43)

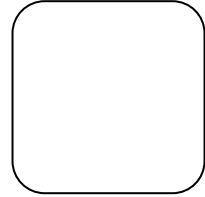
Intermedio: 51 (47-56)

Experto: 63 (58-68)

ANEXO 3

HOJA DE CAPTURA DE DATOS POR OBSERVADOR EXTERNO.

Nombre de Observador:



No. identificación

MODELO CLASICO

	TIEMPO	PRECISION	CALIDAD	CATEGORIA
TRANSFERENCIA				
Observación 1				
Observación 2				
CORTE				
Observación 1				
Observación 2				

MODELO NUEVO

	TIEMPO	PRECISION	CALIDAD	CATEGORIA
TRANSFERENCIA				
CORTE				

REFERENCIAS

1. Derossis AM, Fried GM, Abrahamowicz M, Sigman HH, Barkun JS, Meakins JL. Development of a model for training and evaluation of laparoscopic skills. *Am J Surg* 1998; 175: 482-487.
2. Seymour NE, Gallagher AG, Roman SA, et al. Virtual reality training improves operating room performance. Results of a randomized, double-blinded study. *Ann Surg* 2002; 236: 458- 464.
3. Woodrum DT, Andreatta PB. Construct validity of the LapSim laparoscopic surgical simulator. *Am J Surg.* 2006; 191: 28-32
4. Bridges M, Diamond D. The financial impact of teaching surgical residents in the operating room. *Am J Surg* 1999; 177:28–32.
5. Grantcharov TP, Rosenberg J, Pahle E, Funch-Jensen P. Virtual reality computer simulation. *Surg Endosc* 2001; 15:242– 4.
6. Swanstrom LL, Fried GM. Beta test results of a new system assessing competence in laparoscopic surgery. *J Am Coll Surg* 2006; 202:62–69.
7. Youngblood PL, Srivastava S, Curet M, Heinrichs WL, Dev P, Wren SM. Comparison of training on two laparoscopic simulators and assessment of skills transfer to surgical performance. *J Am Coll Surg* 2005; 200: 546- 551.
8. Chitwood WRJr, Sabiston DCJr. Selected historical perspectives on the evolution of surgical science. En: Troidl H, McKneally MF, Mulder DS, Wechsler AS, McPeck B, Spitzer WO, eds. *Surgical research. Basic principles and clinical practice*. 3rd.ed. New York: Springer-Verlag, 1998; pp. 23-38.
9. Naik VN, Matsumoto ED, Houston PL, et al. Fiberoptic orotracheal intubation on anesthetized patients. Do manipulation skills learned on a simple model transfer into the operating room? *Anesthesiology* 2001; 95: 343- 348
10. Fried GM, Feldman LS, Vassiliou MC, et al. Proving the value of simulation in laparoscopic surgery. *Ann Surg* 2004; 240: 518-528.
11. Peters J, Fried GM, Swanstrom LL, et al. Development and validation of a comprehensive program of education and assessment of the basic fundamentals of laparoscopic surgery. *Surgery* 2004; 135:21–27.
12. Vassiliou MC, Ghitulescu GA, Feldman LS, et al. The MISTELS program to measure technical skill in laparoscopic surgery: evidence for reliability. *Surg Endosc* 2006; 20: 744-747.

13. Brunner WC, Korndorffer JR Jr, Sierra R, et al. Laparoscopic virtual reality training: are 30 repetitions enough? *J Surg Res* 2004; 122: 150-156.
14. Fraser SA, Fried GM, Feldman LS, et al. Setting the pass/fail score for the MISTELS System. *Surg Endosc* 2003; 17:964–967.
15. Powell AC, McAneny D, Hirsch EF. Trends in general surgery workforce data. *Am J Surg* 2004; 188: 1-8.
16. The Society of American Gastrointestinal and Endoscopic Surgeons. <http://www.sages.org> Acceso efectuado el 15 de junio del 2006.
17. Flauhault A, Cadilhac M, Thomas G. Sample size calculation should be performed for design accuracy in diagnostic test studies. *J Clin Epidemiol* 2005; 58: 859-862.
18. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; 33: 159-174.
19. McGinn T, Guyatt G, Cook R, Meade M. Diagnosis. Measuring agreement beyond chance. En: Guyatt G, Rennie D, eds. *Users' guides to the medical literature. A manual for evidence-based clinical practice*. Chicago: AMA Press, 2002; pp.461-470.
20. Kirkwood BR, Sterne JC. *Medical statistics*. 2nd. Ed. Oxford: Blackwell, 2003.
21. Bland MJ, Altman DG. Statistics notes: Cronbach's alpha. *BMJ* 1997; 314: 572.